Development of tools for the analysis of DNA double strand break repair

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in the undergraduate colleges of The Ohio State University

By Jacob M Kroh

The Ohio State University
April 2019

Project advisor: Dr. Ruben Petreaca of the Department of Molecular Genetics

Abstract

Repair of DNA double strand breaks are very important in molecular biology. They play a role in the stability of a genome, and if not repaired quickly and correctly can lead to serious problems for the cell. These problems include genomic instability, chromosomal deletions, duplications, and inversions to name a few. Incorrect repair of these breaks leads to accumulation of DNA damage, so it is important to understand the mechanisms by which these breaks occur and are repaired. Development of new molecular methods is critical to helping researchers determine mechanisms of repair. Here I outline several assays that we have designed for the analysis of *in vivo* DNA double strand break repair.

Acknowledments

I would like to thank Dr. Ruben Petreaca for mentoring and encouraging me to grow as a scientist. I was fortunate enough to be apart of his lab starting in January 2017 and have gained many tools that will aid me in my future as a scientist.

I would also like to offer sincere thanks to the members of my committee: Dr. Ryan Yoder and Dr. John Maharry for their support of, and encouragement to my research. Science is a collaborative effort, and without the support of people from outside my lab I would not have become the scientist that I am.

Finally, to the members of my lab, past and present, I wish to offer my gratitude for the day to day support and company. I have had the pleasure and honor of working with many wonderful young scientists in my time as an undergraduate researcher; these people are the glue that helped hold the lab together through it all.

LIST OF FIGURES

TABLE OF CONTENTS

INTRODUCTION

DNA replication in a heterochromatin context. DNA is a nucleic acid that is responsible for storing information and passing it on to subsequent generations. The DNA polymer is composed of four bases Adenine, Guanine, Cytosine and Thymine and assumes a double stranded helical structure (WATSON AND CRICK 1974). Passage of information from parents to offspring requires duplication of the double helix by a process known as DNA replication. During replication the double stranded- DNA (dsDNA) molecule is split into single strands, copied, and re-annealed to make two helices with the same information, known as the semi-conservative model (MESELSON AND STAHL 1958). The fidelity of the replication machinery is very high to prevent accumulation of mutation during this process. However, mutations may accumulate due to a low rate of spontaneous errors caused by replication.

Another feature of replication is that it proceeds in a semiconservative fashion. This means that each of the new double stranded DNA contains one strand of new DNA and one strand of parental DNA. In addition to spontaneous errors which lead to small base pair changes such as substitutions, deletion and insertions, eukaryotic organisms have another unique problem to deal with during replication. Eukaryotic chromosomes consist of DNA wrapped around proteins called histones. Histones are important in eukaryotic gene regulation as well as packaging the genome in the small nuclear space. Each histone is an octamer (eight protein subunits organized in a complex). Each histones octamer is composed to four sets of two proteins that make up the complex, 2-H2A, 2-H2B, 2-H3 and 2-H4. These histones must be removed so that replication and repair machinery can access the DNA (TABANCAY AND FORSBURG 2006).

Inappropriate removal of these histones can cause the replication machinery to stall and the chromosomes may break resulting in DNA double strand breaks. There are several proteins that aid in the histone removal process with most function during replication, but some have replication independent functions such as chromosome repair and transcription. The functions of these proteins overlap between replication-dependent and replication independent transactions. This thesis focuses on the functions of the replication-independent histone chaperone Hip1 in the yeast model system, Schizosaccharomyces pombe, and the development of assays for the study of smilar systems.

The ability to accurately make mutations is important because it allows reseachers to target a region of a specific gene in a pathway and knock it out precisely. Often it is the case that a genes functions are studied by knocking it out and analyzing the resultant phenotype. However, if the gene is not knocked out precisely, the results could be more ambiguous than desired.

The Hip1 histone chaperone. Hip1 (human HIRA) is a poorly studied non-essential central histone chaperone with replication independent functions (most likely repair). It appears that it makes no contributions to transcription. It overlaps with replication dependent functions by identifying and then moving the histones behind the replication fork (GREENALL *et al.* 2006). Hip1 takes over the replication functions only when the essential replication complex Asf1/Caf1 is compromised. This is important as it keeps the fork from stalling (BLACKWELL *et al.* 2004). As the DNA is replicated, the number of strands effectively doubles, so too must the number of histones. Each histone octamer that is deposited on the newly replicated strands is composed of 50% old and 50% new histones in essence making histone replication also semi-conservative (TABANCAY

AND FORSBURG 2006). This is important to preserve epigenetic histone markers that are required for gene regulation.

In addition to its histone chaperone functions, Hip1 also functions as a transcription factor to up-regulate histone transcription during S-phase. The timing here is important, as up-regulating histone expression is only useful at this stage of replication; so, it makes sense that the protein that moves the histones during replication would also control gene expression (BLACKWELL *et al.* 2004). This function of Hip1 was initially demonstrated by deleting Hip1 and observing histone expression. When Hip1 is deleted, histones are constitutively expressed as opposed to only during S phase in a wild type cell (BLACKWELL *et al.* 2004). Therefore, the function of Hip1 is to repress expression of histones during other stages of the cell cycle. Recent evidence suggests that Hip1 plays a clear role in DNA damage repair as it has been identified to interact both physically and genetically with DNA damage repair genes(GOMEZ *et al.* 2008).

Causes of DNA damage. DNA damage causes changes to its structure. These changes can be caused by several factors which fall into two broad categories: exogenous damage and endogenous damage. Exogenous damage comes from outside the cell and can be of several forms such as chemicals and several types of radiation. Exogenous damage accounts for a small percentage of the total mutations in a genome. It has been shown the majority of an individuals mutations occur before the age of roughly 20 (DEGREGORI 2013). This cannot be explained by exogenous agents, as people under the age of 20 are a protected group not exposed to many chemicals or other forms of DNA damage. Something else must be responsible for this. Endogenous damage explains the fast accumulation of mutations in young individuals. Endogenous mutations (also known as spontaneous damage) are primarily arising, due to errors in replication. It also explains why the protected group of people below age 20 accumulate mutations as quickly as they do as this is the period when their cells are dividing the most. It is worth noting, that endogenous mutations and exogenous mutations differ only in how they are acquired. Both types of damage can cause the same phenotypic changes and are repaired by the same mechanisms.

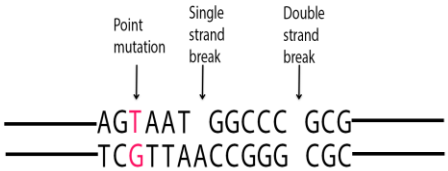


Figure 1. Major types of DNA damage. Three major types of DNA damage can be encountered in a double stranded chromosome. **1**) In point mutation the base on one strand has changed such that it violates the rules of base pairing A-T and G-C. In this case C has changed to T. **2**) In a single strand break, one of the two strands breaks. This can be easily repaired because the information is preserved on the other strand. In this case a T would be inserted opposite A. **3**) A double strand break constitutes severing of both strands. In this diagram the right part of the chromosome could be lost because it completely dissociates from the left part.

DNA damage comes in several forms (**Fig. 1**). Point mutations come in several varieties and are the results of changes of, often, a single base pair. The least serious point mutation is a silent mutation. A change in a codon that leads to the same amino acid is said to have no effect on the phenotype, and thus, is silent. This is possible due the degeneracy of the genetic code. Most of the time however, a point mutation will lead to a change in phenotype. If the change leads to a different amino acid, depending on which one is ultimately produced, this can have anywhere between minimal effects to a complete loss of function. For example, if the amino acid change happens to be in the active site of an enzyme, it could kill the entire function of the enzyme.

Conversely, if the change is in a part of the sequence that is required for the structure of an enzyme it may have almost no effects. It can also happen that the point mutation leads to a premature stop codon. This is a nonsense mutation. Nonsense mutations are often severe as they lead to a gene not being completely translated. Equally severe types of mutations are insertions and deletions (indels). When a base is inserted or deleted, it causes a change in the "reading frame." DNA bases are translated into amino acids in groups of three, (e.g. each group of three codes for a different amino acid) that leads to large amounts of changes down the line from the in/del. This can be suffienct to knockout the gene, and potentially kill the individual.

Another serious type of mutation occurs when the chromosome breaks to some degree. A single strand break (SSB) happens when only one of the two strands is interrupted. A double strand break (DSB) occurs when both stands are interrupted. There are a few ways that a DSB can arise. For example, a conversion of a single strand break to a double strand break can occur due to a collision with a replication fork. This would be considered spontaneous damage. Double strand breaks can also be caused directly by exagenous agents but these events are rare. Usually exogeneous agents only exacerbate the rate at which spontantous breaks may occur. This damage will accumulate over time (RANJHA *et al.* 2018). This would of-course include any genes that happened to be on the lost chromosome.

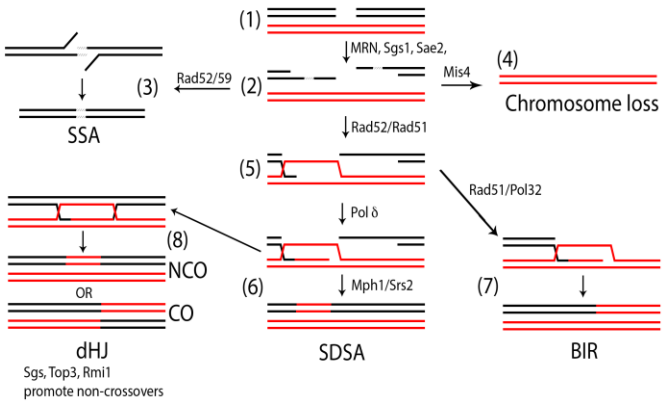


Figure 2. **Pathways of repair of DNA double strand breaks.** A diploid cell with two homologous chromosomes, black and red, sustains a double strand break (DSB) in the black chromosome (1). The DSB is first resected to expose ssDNA required for invasion of donor regions (2). If direct repeats (shaded areas) exist on the same chromosome, the break may be repaired by single strand annealing (SSA) (3). If repair fails, the chromosome may be lost (4). When homology is found elsewhere or on the other homologue (red), the broken ends may invade this region (the donor sequence) (5). In synthesis-dependent strand annealing (SDSA) (6) the invading strand may copy a small region then release and re-anneal. In break-induced replication (BIR) (7) the invading strand may copy to the end of the red chromosome. In this case the right portion of the broken black chromosome is lost. Occasionally a more complex double Holiday Junction (dJH) may be established (8), the resolution of which can result in crossovers (CO) or non-crossovers (NCO). Note that some of these repair outcomes may lead to loss of heterozygosity meaning that the black sequence has been converted to red. If the red sequence contains a recessive non-functional allele, some of these outcomes will convert the functional black allele to the non-functional red allele resulting in complete inactivation of the gene. Some of the genetic requirements for each pathway are indicated.

DNA damage and cancer. Once damage has occurred, it is imperative that it is fixed quickly and correctly. In the case of single and double strand breaks, there are several mechanisms that could be employed to fix this damage. However, there are several problems that much be addressed before any sort of repair can happen. First, the cell must be able to identify the damage. Second, the cell must be able to stop the cell cycle. And third, the cell must select the right repair pathway. In this thesis I will focus only on the type of repair pathways with particular emphasis on DSBs.

Single stranded breaks are easier to repair because the undamaged strand provides a template from which missing DNA can be synthesized. These types of aberrations are more common than double strand breaks and are often easier to repair (CALDECOTT 2008). Even if an incorrect base was added to the break, it would be relatively difficult to fix this type of mutation into the genome, that is to say that it becomes a part of the genome without the ability to determine where, or even if, a mistake had been made. After fixation, the mutation is said to be transmissible or passed to progeny. It would take two rounds of replication to fix such a mutation into the genome. After the first round of replication, there would still be evidence of the mutation that could still be repaired because only one strand of the DNA will have the incorrect base that will mismatch with the base on the undamaged strand. The cell has good mechanisms to detect such mismatches and efficiently repair them. However, after the second round of replication, the mutation can no longer be detected, and thus is said to be fixed into the genome.

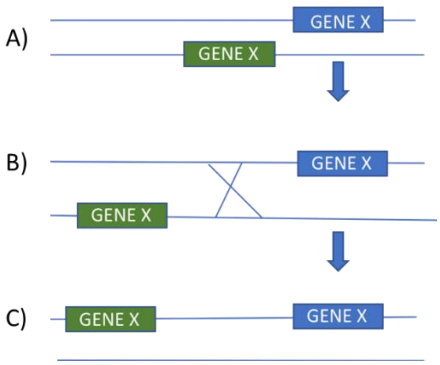


Figure 3: **Example of potential CNV mechanism.** **A**) A homologous chromosome is misaligned such that the homology is not between genes, rather it is found in the segments before and after the genes respectively. There is no real homology between the two chromosomes but because only 3-5 base pairs may be required the cell is "tricked" into finding homology where there shouldn't be any. **B**) This leads to homologous recombination, not between true homologous regions or alleles of the same gen as expected, but in the intervening segments. Once a repair structure has been established, the sequence in green copies off the sequence in blue. **C**) Finally, we have the chromosome that has now acquired a second copy of the gene x. Thus, its number of copies of the gene x has increased. Note also that this leads to a translocation from the blue chromosome to the green chromosome. These mechanisms, and others can also lead to deletions as well.

For a double strand break however, the problems are magnified and are the most serious type of DNA damage. This is because when both strands of the helix are lost it is harder to recover the missing sequence. The only choice is to copy from another part of the genome with similar sequence (homologous). If improperly repaired it would lead loss of genetic information or alteration which in many cases can be lethal.

There are many ways that DNA DSBs can be repaired (**Fig. 2**). Some of the pathways of repair are more error-prone while others are more error-proof in nature. Because of how deletarious double strand breaks are, it is important that multiple pathways exist for their repair (MEHTA AND HABER 2014). Two of the most common methods are end-joining (not shown) and homologous recombination (HR). These methods differ greatly in speed and how accurately they repair damage. They are also dependent on the time of the cell stage with end joining occuring primarily in G1 and HR in S-phase and G2.

End-joining (not shown in the figure) is a quick way of fixing a DSB. As the name would imply, end-joining is a process where the cell takes a DSB and ligates it back together. There are broadly two types of end-joining: non-homologous end joining (NHEJ), and microhomology mediated end-joining (MMEJ) (SALLMYR AND TOMKINSON 2018). End-joining only works on two-ended breaks, as it wouldn't be possible to ligate a one-ended breaks (discussed later). This mechanism also occurs primarily in G1 because it does not require a sister chromatid to find homology. Endjoing is highly mutagenic, because it leads to some alteration of the

sequence. This is due to some degradation of the sequence at the two ends before they are ligated together (SALLMYR AND TOMKINSON 2018). Being cell cycle independent and not requiring a template makes this type of repair fast however, thoughquite mutagenic. However this type of repair is prefered in higher eukaryotes including humans because 95% of their genomes are not transcribed decreasing the posibility of mutations altering gene function.

Another repair mechanism is *Break Induced Replication* (BIR). For this to happen, the ends of the double stranded break must be processed such that a single stranded region is present. The single stranded section of the chromosome is bound by RPA which protects it from degradation. This process requires RAD51 which binds the single strand DNA. Then a homology search begins. Once complete, the strand of ssDNA invades the homologous region begins replication to the end of the chromosome. The problem here is that this is a highly mutagenic process as it leads to an asymmetric crossover from the point of replication to the end of the chromosome (SALLMYR AND TOMKINSON 2018).

Another mechanism of two-ended break repair is *Single Strand Annealing* (SSA). This type of repair requires that there are repetitive elemens on either side of the break. After the break, various exonucleases chew up one of the strand leaving uneven ends of single stranded DNA. Because the DNA sequences are repeated on either side of the break, they are in essence identical sequences so the two strand of ssDNA anneal together and are ligated to repair the break. This type of repair is highly mutagenic because the sequence, along with any genes that happened to be in between the repetitive elements, is deleted. This can be a serious problem if the intervening gene was an essential gene as it could kill the cell. An even more serious problem can arise from deleting a tumor supressor gene as the cell will lose its ability to control cell división and may become a cancer precursor.

Synthesis dependent strand annealing (SDSA) is the major mechanism of repair of mitotic breaks. In this mechanism, the missing sequence from the broken chromosme is copied from a homologous región elsewhere, tipically the homologous chromosome. Note that SDSA can lead to conversión of the black allele to the red allele (**Fig.2**). This is known as loss of heterozygosity because half of the information (black) has been lost. Other mechanisms such as *Double Holiday Junctions* (dHJ) are rare and restricted mainly to meiosis.

If repeats are not found on either side of the break the cell will look for homology elsewhere. Perhaps the biggest problem is where to find homology. It is not at all obvious where this may be, but it is extremely important to find something similar to the original so that the fidelity of the sequence is preserved. However, this is not always practical. Sometime better to find something that is near by to repair the damage quickly. It is for this reason that the definition of homology can be loosened to allow for a quicker resolution to the search, and the damage addressed (HASTINGS *et al.* 2009). One mechanism similar to end-joining known as microhomology mediated repair appears to also be favored in human cells. It was observed by the Hastings group, that at times the number of similar base pairs required in the homology search can be as few a three (HASTINGS *et al.* 2009). This can be a convienient way of dealing with damage as homology could be found virtually anywhere. This method of microhomology search also would be very useful for potentially fixing damage quickly. However, this can lead to the problem of misidentified homology where different chromosomal regions are fused with other regions. Some other problems could occur if homology was found on the same chromosome as it can cause deletions.

This type of microhomology repair can lead to copy number variation (CNV), a common occurrence in cancer cells (VISSERS *et al.* 2009). CNV refers to genomes with either a greater or fewer number of genetic regions than the normal two copies seen in wild type cells. It is worthy of note that CNV is an important factor in evolution and individuals are often able to deal with

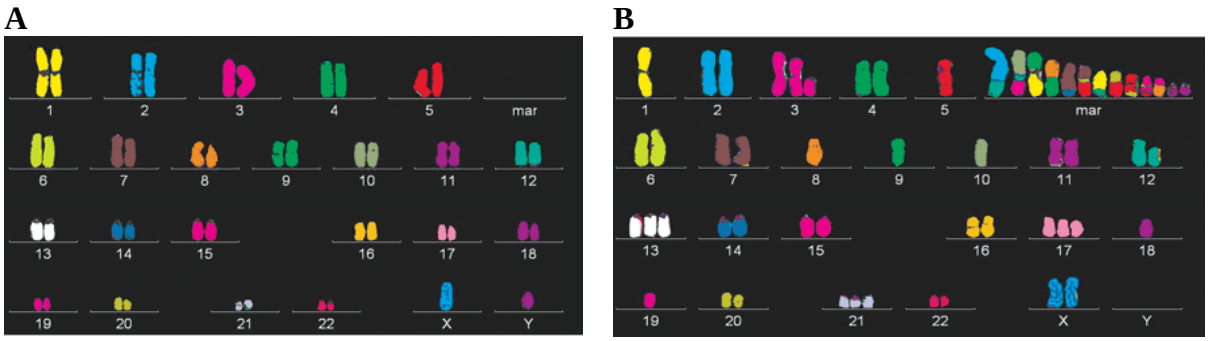


Figure 4. Microscopic chromosome karyotypes. A. Each of the 22 chromosome pairs and the sex XY chromosomes are painted with a different color to be distinguished from each other under the microscope. **B.** A cancer cell has reorganized its chromosomes such that many can no longer be classified as a normal 1 to 22 chromosomes. It is said that this cancer cell has been transformed.

high levels of CNV in certain cases (**Fig. 3**). For other cases, certain genes are highly sensitive to these changes, which could lead to serious genetic problems for the individual. It arises due to duplication or deletion events and is generally either a short segment repeated, or a longer segment repeated. It is also possible for entire genes to be duplicated. It is thought that this could be the result of either non-homologous end joining, or micro homology mediated end joining events. Such is the case for the alpha amylase gene studied by the Perry group (PERRY *et al.* 2007). But the most important consequence of this type of repair may be a vast amount of chromosomal instability (CNVs) seen in cancer cells (**Fig. 4**).

Another, often more serious problem occurs when homology is found on different chromosomes. This can lead to the most error-prone of HR pathways, resolution of a double Holliday junction leading to a mitotic crossover. This is often termed a translocation. Crossovers are common in meiosis, and contribute to genetic variation when in meiosis. However, the mechanisms originated as a response to DNA damage. When a translocation via mitotic crossover occurs, it can drastically effect the genomes stability. This is reflected in the infamous "Philadelphia chromosome "where ch22 incurs damage and is repaired via homology found on ch9. The result is that two wholy different chromosomes have been created. It so happens that the break occurs in the middle of a key cell cycle regulating gene, the now inactive gene is unable to regulate the cell cycle. This leads to uncontrollable cell division, a hallmark of cancer.

RESULTS

Molecular methods are important as a means of assessing the potential effects of DNA damage as well as the mechanisms of repair. Different assays can be used to assess different pathways that the cell employs for DNA damage repair. My project focused on using existing *in vivo* DSB repair assays to study break repair. Additionally, I developed another assay to introduce genetic changes *in vitro*.

Use of existing assays to study DSB repair. An early Ade-His-Ade assay designed designed by Osman et al (OSMAN AND WHITBY 2009). employs a construct where two non-functional

heteroalleles of the *ade6* gene flank a functional *his3+* gene (**Fig 5A**). This makes the cell an auxotrophic for adenine, and prototrophic for histidine. This means that these cells can be plated on media lacking histidine, and they will grow. If grown on media lacking adenine, they die. When a DSB break occurs between the two *ade6* alleles, the cell will attempt to repair by several

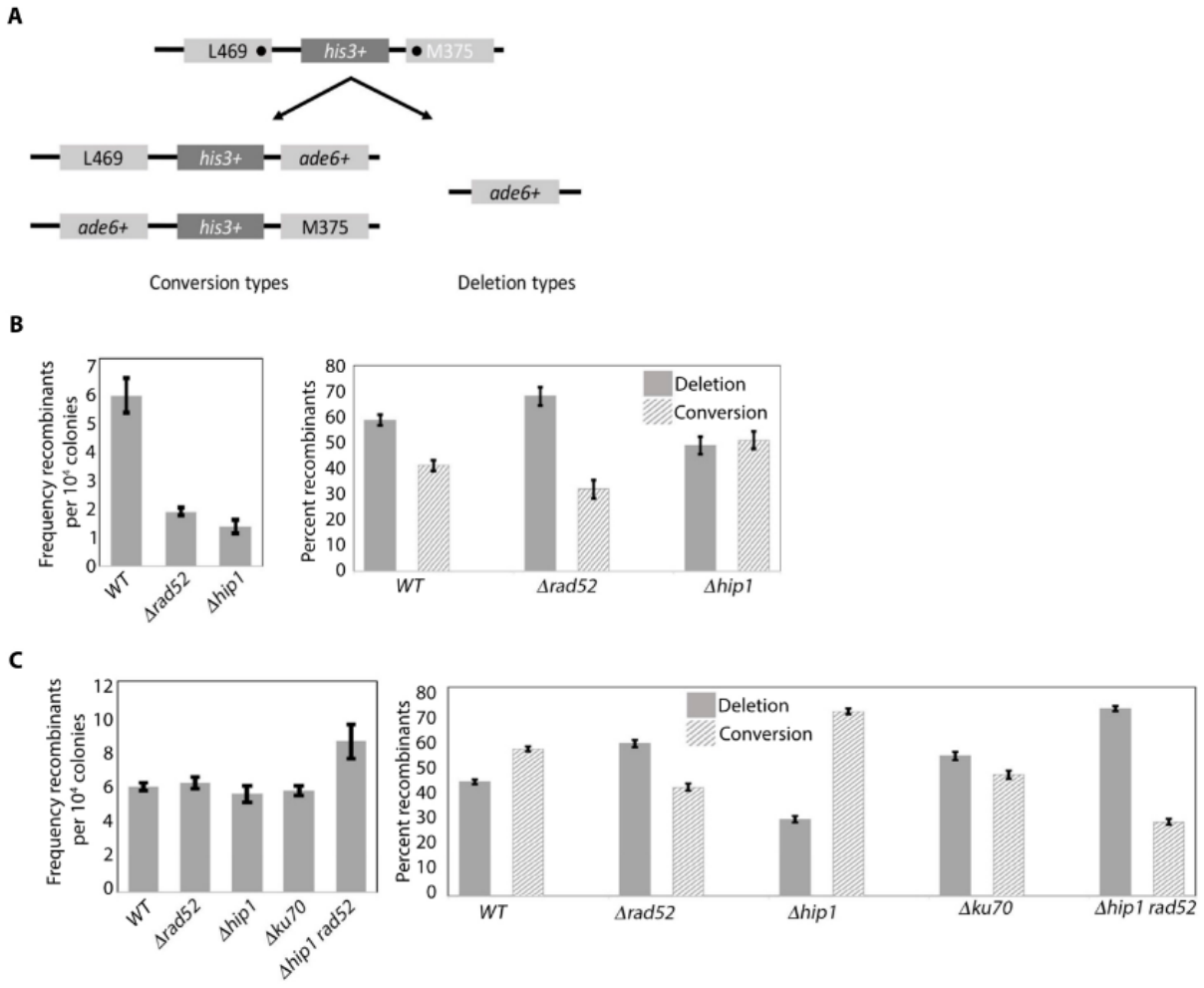


Figure 5. *hip1+* biases recombination towards deletion at regions of tandem repeats. A. A previously described genetic assay to study recombination outcomes at regions of tandem repeats (AHN *et al.* 2005). Two loss of function *ade6* alleles, *ade6-L469* and *ade6-M375* flank the functional *his3+* gene. The locations of the point mutations of the ade6 are indicated by the black dot. is A random break between the two *ade6* alleles could be repaired by two outcomes, conversion and deletion. Deletion outcomes lose the functional *his3+* gene. **B**. Left panel: Deletion of *rad52* and *hip1* show a decrease in frequency of recombination which is assayed by total colonies appearing on EMM-Ade plates. Right panel: Data pulled by percentage of deletion (Ade+His-) and conversion (Ade+His+). The assay was performed as previously described (OSMAN *et al.* 1996; OSMAN *et al.* 2000; AHN *et al.* 2005). **C**. Increasing the time in non-selective media changes the recombination outcomes. Cells were released in non-selective YES at 100cells/microliter for 48hs then plated on selective EMM-Ade and subsequently replica plated on EMM-Ade-His. Left panel: The frequency of colonies appearing on EMM-Ade plates is similar for all strains tested. Right panel: When pulled by percentage of deletion and conversion, loss of *hip1* decreases deletion outcomes while loss of *rad52* decreases conversion outcomes. For all graphs error bars show standard error.

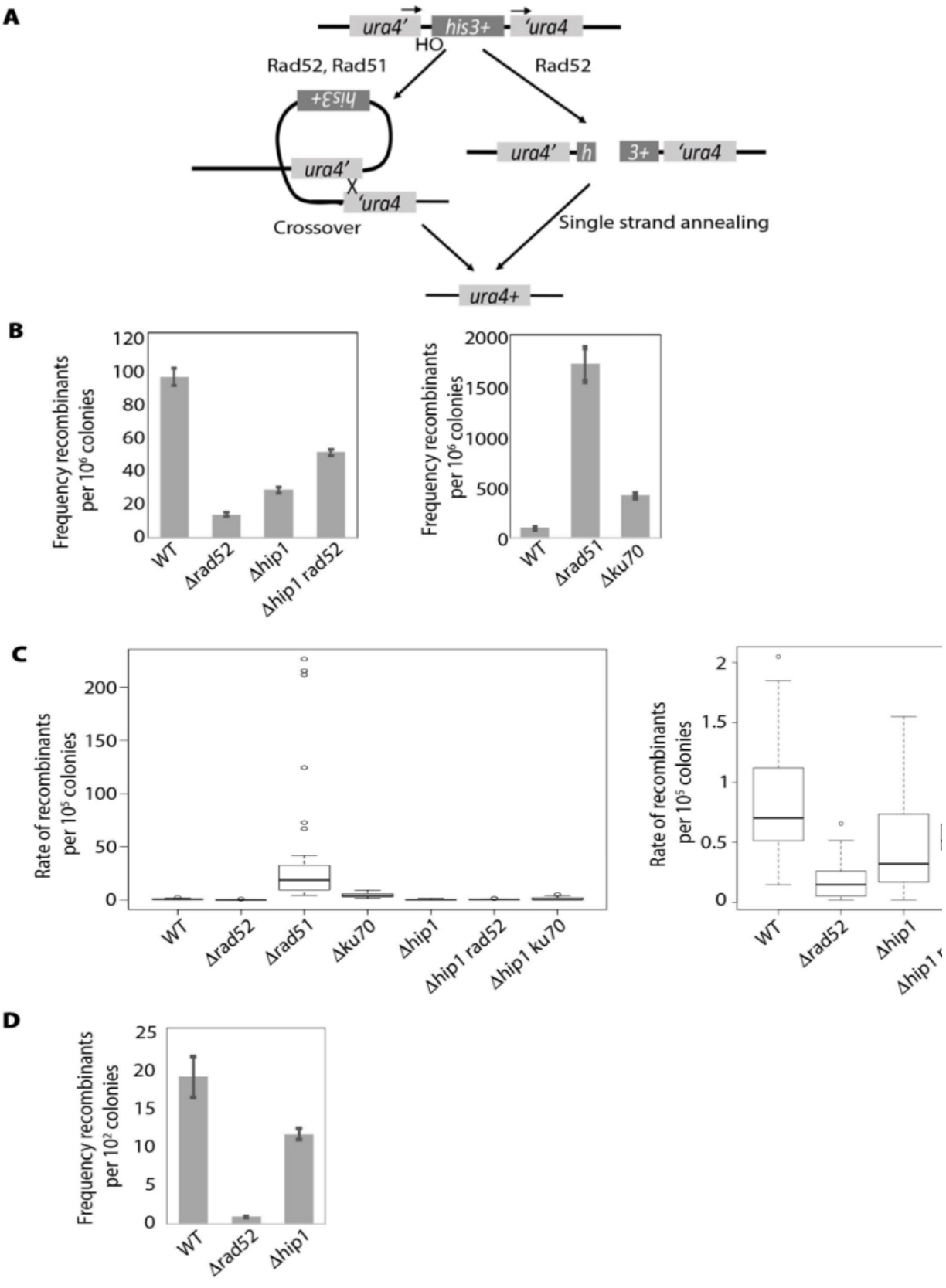

A
ura4'
his3+
'ura4
HO
Rad52, Rad51
Rad52
his3+
ura4'
X
'ura4
Crossover
ura4' h 3+ 'ura4
Single strand annealing
ura4+
B
Frequency recombinants per 10⁶ colonies
120
100
80
60
40
20
0
WT
Δrad52
Δhip1
Δhip1 rad52
Frequency recombinants per 10⁶ colonies
2000
1500
1000
500
0
WT
Δrad51
Δku70
C
Rate of recombinants per 10⁵ colonies
200
150
100
50
0
WT
Δrad52
Δrad51
Δku70
Δhip1
Δhip1 rad52
Δhip1 ku70
Rate of recombinants per 10⁵ colonies
2
1.5
1
0.5
0
WT
Δrad52
Δhip1
D
Frequency recombinants per 10⁷ colonies
25
20
15
10
5
0
WT
Δrad52
Δhip1

Figure 6. hip1+ promotes repair by SSA at regions of tandem repeats. A. An assay to study repair of breaks by single strand annealing (SSA)(LI *et al.* 2013). Two *ura4* alleles that have 300bps of identity (arrows) flank the functional *his3+* gene. The *S. cerevisiae* homothallic endonuclease (HO) was cloned between the left *ura4'* allele and *his3+.* Repair of breaks arising between the two *ura4* alleles is hypothesized to occur by a crossover which relies on both Rad52 and Rad51 or by single strand annealing which relies only on Rad52. **B**. The *ura-his-ura* assay is as a reporter for SSA. The frequency of recombinants is drastically decreased in cells lacking *rad52* and *hip1* but not in *rad51* and *ku70*. This suggests that repair occurs primarily through the SSA pathway and Hip1 makes contributions to this pathway. Error bars show standard error. **C**. The rate of recombination (frequency per generation) is decreased in cells lacking *rad52* and *hip1* but not affected by *rad51*. the data was analyzed using the Lea-Coulson method (http://www.keshavsingh.org/protocols/FALCOR.html) and graphs were generated with the R statistical program. **D**. Recombination frequency of induced breaks. Breaks were induced by activating the HO endonuclease after de-repressing the *nmt1* promoter. Shown is the frequency of EMM-Ura colonies with standard errors.

pathways. One possible means of repair is single strand annealing (SSA). SSA will result in deletion of the *his3+* gene and reconstitution of the functional *ade6*. Note that the positions of the mutations in the ade6 heteroallele allows deletion of non-functional *ade6* regions to produce a WT copy of the gene. This is because SSA found homology between the two *ade6* fragments which function as tandem repeats and deleted the histidine gene in the process. The phenotype of such an event will result in cells that can now grow on adenine but not histidine. This assay also allows repair by SDSA which will result in restoration of *ade6* function without loss of the *his3* function. Thus, this assay allows for quantification of deletion.

Our laboratory designed a Ura-his-Ura assay similar to the Ade-His-Ade assay (LI *et al.* 2013). This particular assay tests mainly for single strand annealing (SSA) (**Fig.6A**). This is because the ura4 heteroalleles are truncations with overlapping 200bp regions of homology and SDSA is not possible. This overlap is sufficient for homology search required for SSA when a break occurs. This assay shows how homology can be found on the same chromosome and how SSA can lead to interstitial deletions.

The role of hip1 in DSB repair. Novel interactions between Rad52, Hip1 and another hisone acetyltrasnferase, Mst1, were identified in a two-hybrid screen (GOMEZ *et al.* 2005; GOMEZ *et al.* 2008). Here we investigated how Hip1 affects repair of spontaneous DNA breaks arising during replication.

We employed the genetic assays described above. Deletion of *hip1* has been previously shown to be sensitive to DNA damage drugs (*Anderson et al., 2009*). We also find that when grown on media containing methyl methanesulfonate (MMS), an alkylating agent or camptothecin (CPT), a topoisomerase inhibitor, cells lacking the *hip1+* display marked sensitivity to these drugs (**Fig.7A**). *hip1Δ* is not as sensitive as of *rad52Δ*. Nevertheless, these data indicate that *hip1+* deletion affects some form of DNA damage repair. Hip1 has been previously shown to be required for establishment of a functional centromere (BLACKWELL *et al.* 2004). Using an assay to look for gene conversion (CULLEN *et al.* 2007) we found that indeed deletion of *hip1+* results in higher level of chromosome loss (Ade- His) than WT and when compared to chromosome loss seen in rad52Δ (**Fig. 7B**). We then used the assay to look for gene conversion of the ade6-M216 allele which is closest to the centromere arising from repair of spontaneous breaks. We found that a statistically significant subset of outcomes shows only gene conversion of this locus (Ade-His+). An Ade- His+ phenotype could result by two different pathways: 1) gene conversion of *ade6-M216* to *ade6-M210*, or 2) possibly deletion of the *ade6-*

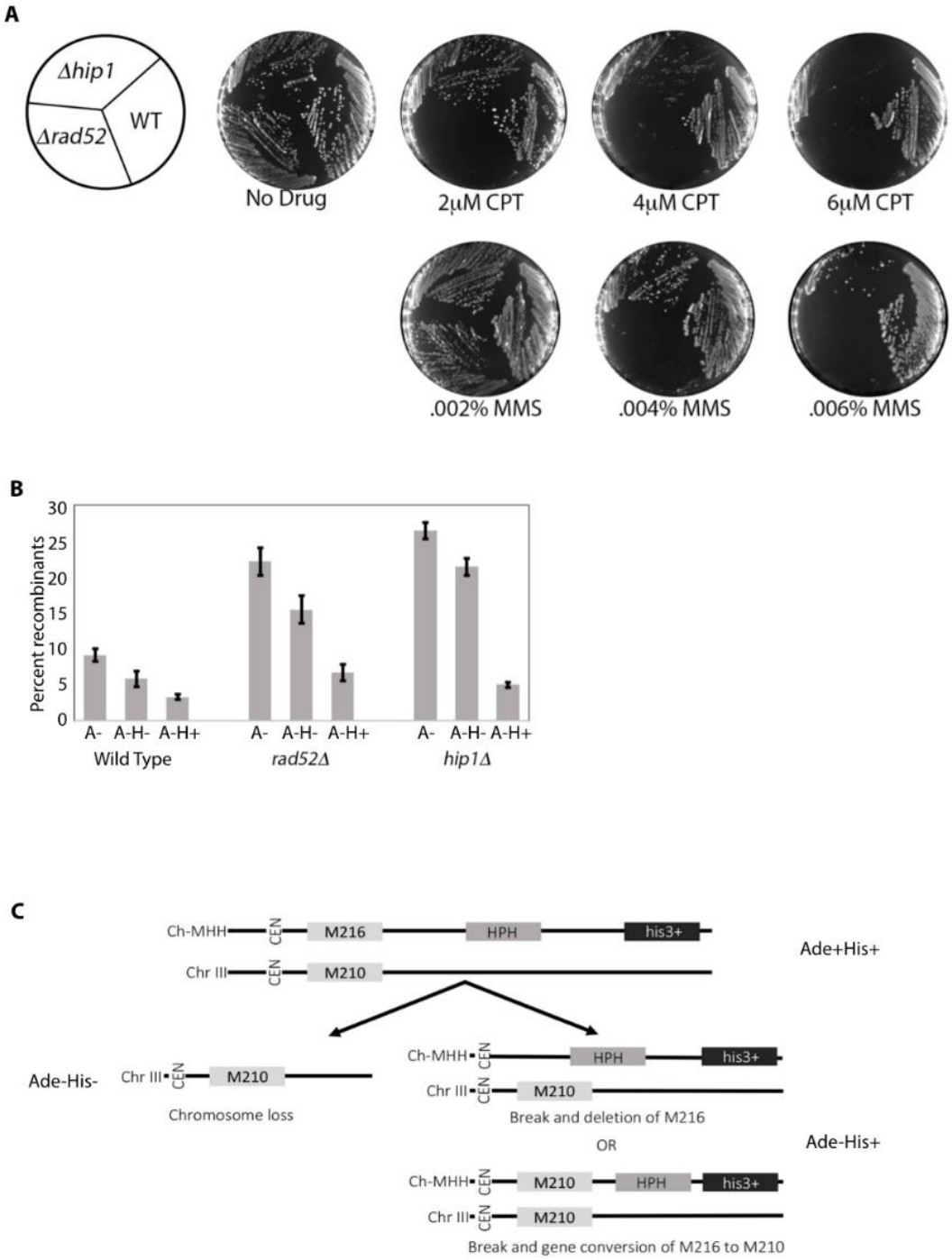


Figure 7. Hip1 inhibits gene conversion next to centromere. A. *hip1Δ* and *rad52Δ* are sensitive to DNA damage. *hip1Δ* and *rad52Δ* strains were struck out on YES or YES plus the indicated drugs. Plates were incubated at 32C for 4 days. **B**. We used a previously described assay to look for gene conversion next to centromere. The assay was performed as indicated in the Prudden et al paper (PRUDDEN *et al.* 2003). Total Ade- colonies represents all colonies that grew on EMM-Ade. Ade-His- colonies were calculated as (EMM-Ade colonies)-(EMM-His colonies). Ade-His+ colonies were calculated as (EMM-Ade colonies)-(Ade-His- colonies). Graphs show the mean and standard error. **C**. Schematic diagram showing the outcomes measured using this assay. The Ade- colonies could be either Ade-His- or Ade-His+.

M216 allele (**Fig.7C**) We could not differentiate between these two possibilities using this assay. Nevertheless, we find that these events are reduced and different from the other events when *hip1+* is deleted. We hypothesized that hip1 might affect some form of DNA double strand break repair.

Hip1 biases recombination towards deletion. We wanted to investigate in more detail the function of hip1+ in repair of DSBs so we turned to a more specific assay where a tandem duplication of the ade6+ gene is separated by the *his3+* transcript (AHN *et al.* 2005). When repair outcomes of *ade6* are analyzed they fall into two categories: deletion and conversion (**Fig.5A**). A decrease in the spontaneous frequency of recombination was seen in both *rad52Δ* and *hip1Δ* suggesting that both genes affect DSB repair (**Fig. 5B**). Repair of spontaneous breaks in WT cells should occur about 65% by deletion and 35% by conversion. When we repeated the experiment as previously described (*Osman et al., 2000*) we found that we did get about this ratio (**Fig. 5B**). We also saw an increase in deletion events in *rad52Δ* as has been previously reported(OSMAN *et al.* 2000; AHN *et al.* 2005). However, *hip1Δ* increased conversion events and decreased deletion events when compared to WT. This suggested that loss of *hip1+* might affect deletion events.

We hypothesized that longer growth times in non-selective media may allow cells to choose different repair pathways, particularly the ones requiring more complex re-arrangements. We decided to alter the assay by releasing cells in liquid YES at 100cells/uL then allowing them to grow for 48h after which cells are counted then plated on selective media. We found that under these conditions the frequency of ade6+ repair is not significant in rad52Δ and Hip1 Δ when compared to WT similarly to what has been previously reported (**Fig. 5C**). The *hipΔrad52Δ* double mutant showed a slight increase compared to WT but not statistically significant from *rad52Δ*. When we analyzed deletion and conversion outcomes, we found that the wild type strains showed an increase in conversion outcomes consistent with the hypothesis that allowing more time might bias repair through the more complex and conservative pathways. *rad52Δ* strains still showed an increase in deletion outcomes indicating that it is required for the conservative pathways. Remarkably *hip1Δ* showed an increase in conversion outcomes (**Fig. 5C**). Because conversion outcomes require more complex forms of repair this may indicate that allowing more repair time allows strains lacking *hip1+* to choose the more complex pathway. These data also suggest that *hip1+* promotes deletion outcomes.

Hip1 promotes repair by single strand annealing (SSA). We previously reported an assay that monitored chromosomal instability (LI *et al.* 2013) but here we show that this assay is a reporter specific for deletions (**Fig. 6A**). We used this assay to monitor repair of spontaneous breaks. We find that rad52Δ shows a statistically significant decrease in frequency of repair while remarkably *rad51Δ* shows an increase in repair when compared to WT (**Fig. 6B**). These data mean that most repair occurs through the SSA pathway and is inhibited by loss of *rad52+. rad51+* inhibits SSA as it presumably biases repair through other homologous recombination pathways and deletion of *rad51+* allows more repair outcomes through the SSA pathway.

Although SSA relies on Rad52, in higher eukaryotes and fission yeast, Rad52 is not essential for all forms of homologous recombination repair (YAMAGUCHI-IWAI *et al.* 1998; OSMAN *et al.* 2000). The fact that some repair still occurs in the absence of *rad52+* indicates either that at a low percentage *ura4+* may be reconstituted by some other form of repair that is not SSA. Note

also that deletion of *yku70+* does not affect the repair suggesting that NHEJ does not contribute to restoration of the functional *ura4+* gene. This assay is not unlike others reported before but has the benefit of not having to monitor outcomes by sequencing or gel electrophoresis. In addition, we believe that because the two tandem *ura4+* repeats are so close to each other there is a space constraint for the cell to do a crossover and most repairs occur by SSA (**Fig. 6A**).

Deletion of *hip1+* drastically reduces the repair outcomes suggesting that *hip1+* also makes contributions to the SSA pathway although it does not appear to be as important as *rad52+. hip1Δ* suppresses the *rad52Δ* phenotype (compare *rad52Δ* with *rad52Δhip1Δ*) indicating that *hip1+* inhibits some other form of repair that does not rely on *rad52+.* Taken together, these data suggest that this assay can be used to study mainly repair by SSA and that *rad52+* and *hip1+* function in the SSA pathway to promote intra-chromosomal deletions at regions of tandem repeats.

Recombination rate is decreased in cells lacking hip1. Because we investigated spontaneous breaks arising from replication stress, we also assayed the rate of recombination (frequency per generation) (**Fig. 6C**). Indeed, we see the same pattern, where both *rad52Δ* and *hip1Δ* show a decrease in rate while *rad51Δ* and *yku70Δ* show an increase. We present these data as boxplots (with whiskers) to show quartiles and outliers. Clearly the variation in data for *rad51Δ* is distinctly higher than all the other mutants even though is somewhat noisy. Both *rad52Δ* and *hip1Δ* show a lower rate of recombination than WT. While it is clear that the *rad52Δ hip1Δ* suppresses defects in SSA, the repair rate is not restored to WT levels suggesting that although other repair mechanisms may play a role, SSA is the predominant repair pathway.

This assay can also be used to study induced breaks (**Fig. 6D**). We placed the yeast homothallic HO endonuclease between the left *ura4+* repeat and *his3+.* Upon induction of a break we see a marked decrease in repair in *rad52Δ* consistent with the observed results for spontaneous breaks. *hip1Δ* also decreases the frequency of repair from induced breaks but does not appear to play as much a role as in repair of spontaneous breaks.

An assay to make genetic changes in vitro. The SIMPLE assay developed in conjunction with this lab and a collaborator, Hovik Gasparyan, is a way of site directed mutagenesis using restriction enzymes (GASPARYAN *et al.* 2018). While multiple different methods have been developed for targeted site-specific mutagenesis, deletions, and small insertions of plasmid DNA, no single currently available approach is able to produce all of these changes in a single PCR reaction (TESSIER AND THOMAS 1993; LING AND ROBINSON 1997; ZHENG *et al.* 2004). In our method, we took advantage of the unique characteristics of a pair of Type II restriction enzymes, SapI and AarI (GRIGAITE *et al.* 2002). Both of these enzymes have several properties which make them conducive to the SIMPLE method (summarized in **Fig. 8A**). First, both have recognition sequences of 7 bps, making their cut sites relatively infrequent in commonly used cloning vectors. Second, they cut on the 3′ side of their recognition site rather than the position where they bind. Third and most importantly, there are no constrains for the sequence being cut (Ns), allowing an overhang of any sequence to be generated.

We reasoned that by using these properties of SapI/AarI, we would be able to edit specific nucleotides in any plasmid via PCR amplification, restriction enzyme digestion, and a final ligation. As an example, we show a partial hypothetical DNA sequence on a plasmid, with a single nucleotide that will be targeted for mutagenesis highlighted in red (**Fig. 8B**). The primers would be designed to contain a 3- to 4- nucleotide spacer sequence at their 5′ end, followed by the SapI

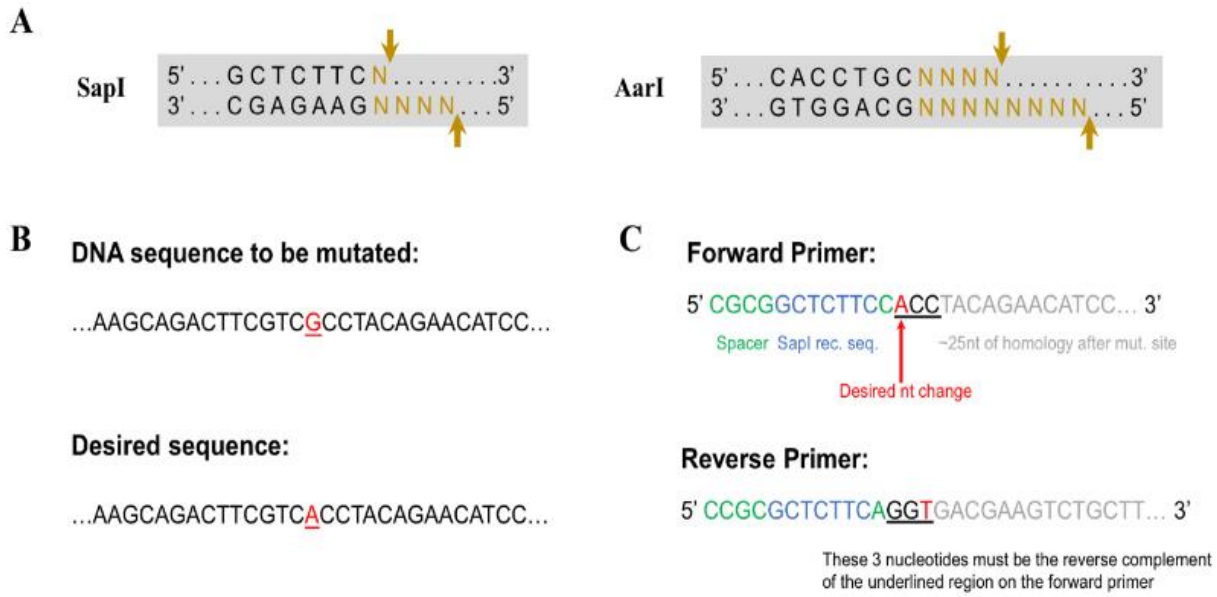


Figure 8. Primer design for introducing point mutations using AarI and SapI. **A**. Recognition sites and cutting patterns for the SapI and AarI restriction enzymes. **B**. Hypothetical sequence that will be targeted for mutagenesis. **C.** Primers that can be used to mutate the sequence shown in panel B with the SIMPLE method, using the SapI enzyme.

(or AarI) recognition site, then 1 (for SapI) or 4 (for AarI) nucleotides of spacer sequence. Directly downstream of this spacer, there are up to 3 (for SapI) or 4 (for AarI) nucleotides that can be engineered to contain any sequence. Lastly, the remaining ~25 nucleotides will be used for targeting the primers to the appropriate sites for PCR amplification. The specific steps for designing primers are outlined in **Fig. 8C**. PCR amplification of a plasmid using our primer design scheme would lead to production of a linearized, full length version of the plasmid, flanked by the desired point mutation and the SapI/AarI sites on both the 5' and 3'
ends. This PCR product can then be digested with the appropriate enzyme, generating a pair of compatible overhangs, which when ligated together would reconstitute the original plasmid containing the desired mutation (**Fig. 9A**). Similarly, this approach can be modified to delete any part of a plasmid without the need for specific restriction enzyme sites, generating a perfect sequence without any "scars" left from the cutting and ligation. To perform a deletion, the PCR primers will be designed in the same way as above, but made complementary to the sequences flanking the desired deletion site. The PCR, digestion, and ligation steps will be identical to those used for mutagenesis (**Fig. 9B**).

As a proof of principle, we decided to use the SIMPLE method to make two separate point mutations (A46 → G and C1466 → T) and a partial deletion (601–1101) of the S. pombe ade6+ ORF using both SapI and AarI. A short sequence of the relevant region for the C1466 → T mutation is shown in **Fig. 10A**. The A46 → G change is not shown. To determine the efficiency of our

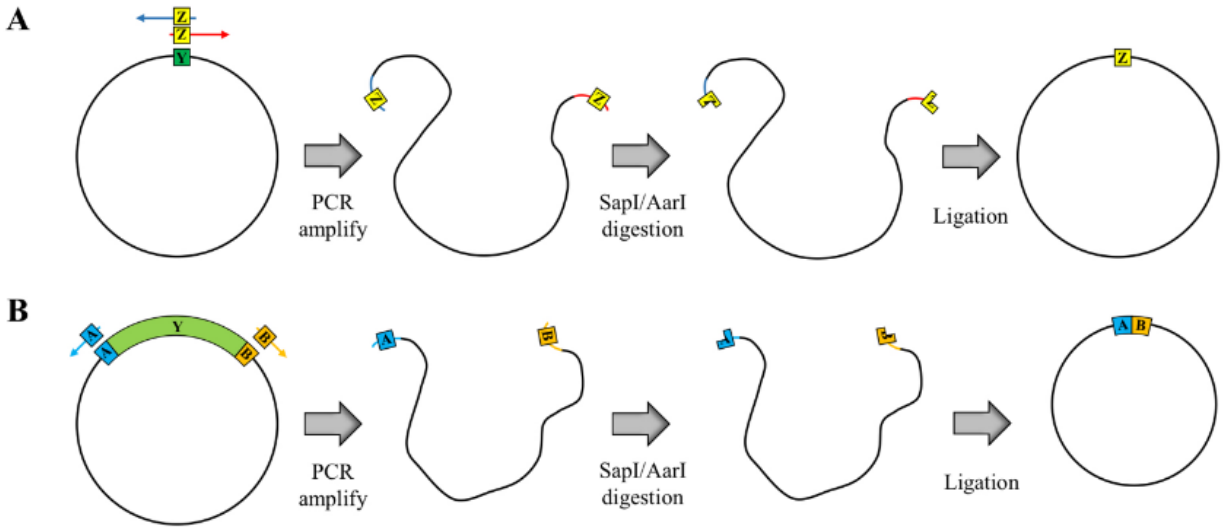


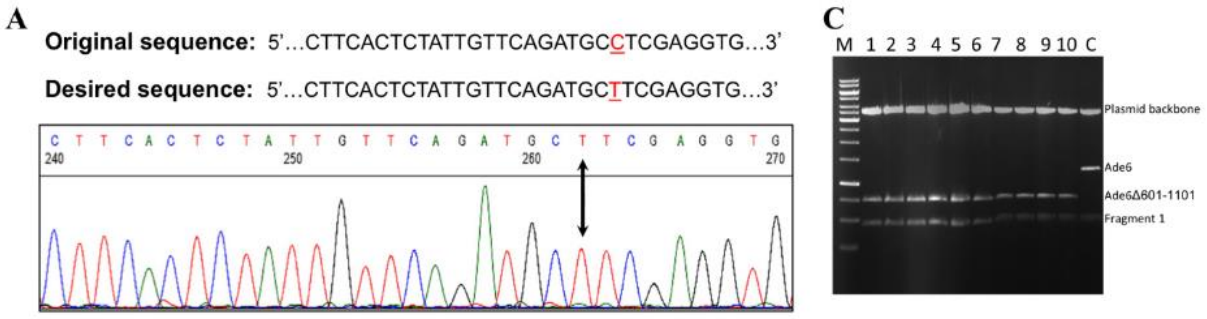
Figure 9. Schematic diagram showing how the SIMPLE method can be used to perform site-directed mutagenesis (A)and a deletion within a plasmid (B). The desired sequence to be targeted for either mutagenesis or deletion is labeled with a Y in both figures, while the desired nucleotide change is labeled with a Z, and sequences flanking the deletion site are labeled with **A** and **B**.

method, we sequenced five possible clones from each of the mutations and deletion, and found that our method had a success rate of 80% (**Fig. 10B**). We also screened several potential clones for the deletion by restriction enzyme digestion followed by gel electrophoresis (**Fig. 10C**).

The SIMPLE method can also be adapted to make small insertions in a single PCR step. It is commonly desirable to clone small epitopes such as the HA, MYC, or FLAG to tag gene cloned in a plasmid. However, researchers often run into the issue of having no convenient restriction site at their desired position, or not being able to clone their tag in the proper reading frame. The SIMPLE method has the ability to circumvent all of these issues, as there are no restrictions to where the cut sites can be located. As a proof of concept, we inserted the FLAG epitope into a precisely defined position within the *ade6+* ORF. As before, the primers were designed to contain a spacer, followed by the enzyme recognition site, an additional spacer, then half of the FLAG epitope (per primer), and finally the targeting homology sequence. After the PCR amplification, digestion, and ligation, the two compatible ends generate the FLAG tag in the proper reading frame.

The unique properties of SapI and AarI can also be utilized to ligate together multiple large DNA frag- ments in any desired order. Each fragment can be amplified via primers that, when digested, will contain a specific overhang. By generating pairs of compatible overhangs, multiple fragments can be “stitched” together to construct a plasmid. In theory, this approach can be used to produce 4 3 = 64 unique overhangs using SapI, and 4 4 = 256 unique overhangs using AarI. However, further work needs to be done to assess the practical feasibility of this concept.

Lastly, we propose that additional enzymes with similar properties can also be utilized to perform the SIMPLE method. We have chosen to use SapI and AarI for this study as their



B

Mutation/ Deletion	Enzyme used	Result
ADE6 A46→G	SapI	4 out of 5 sequenced plasmids worked
	AarI	4 out of 5 sequenced plasmids worked
ADE6 C1466→T	SapI	3 out of 5 sequenced plasmids worked
	AarI	5 out of 5 sequenced plasmids worked
ADE6 Δ601-1101	SapI	4 out of 5 sequenced plasmids worked
	AarI	4 out of 5 sequenced plasmids worked

Figure 10. Proof of principle for the SIMPLE method. A. Sanger sequencing chromatograph showing a directed site- specific mutation in the S. pombe ade6+ gene where the C 1466 was changed to a T. The A46 → G change is not shown. **B**. Summary of all attempted mutations and deletions using the SIMPLE method. For all methods, we screened five plasmids. **C**. Agarose gel showing the targeted 500-bp deletion of the ade6+ ORF. M, size markers; 1–10, screened deletions; C, control plasmid. The plasmid was cut with EcoRI which cuts three times: on either side of the ade6+ gene (releasing the ade6 fragment) and one more time (releasing Fragment 1 and Plasmid backbone). We screened 10 candidates with the deletion. Note the shorter *ade6* fragment.

recognition sequence of 7 bps is relatively rare, and does not occur in our plasmids. However, for researchers who do plan to utilize our approach, we have assembled a list of additional enzymes that have a recognition se- quence of at least 6 bps, and cut distal to their binding sites on the 3' side. Theoretically, any of these enzymes should also be compatible with SIMPLE cloning.

METHODS

Recombination assays. *Ch-MHH recombination assay*. Cells were struck out from -70°C freezer onto YES and incubated at 32°C until colonies appeared. The YES plates were then replica-plated onto EMM-Ade (and YES master) to ensure that the minichromosome is still present. Colonies that grew on EMM-Ade were picked from the YES plate and cells were counted. For each count, 500 cells were spread onto two different YES plates. The YES plates were incubated at 32°C for 6 days to allow for spontaneous recombination. After 6 days, colonies on all plates were counted then the plates were replica plated onto EMM-Ade, EMM-His, EMM+HPH. The dropout plates were placed back at 32°C for 4 days after which the colonies on these plates were also counted.

ade6-L469/his3/ade6-M375 recombination assay. We followed the procedure as previously described (OSMAN *et al.* 2000) except that we used Edinburgh minimal media (Sunrise Science). Strains were first struck onto YES from the -70°C freezer and allowed to grow at 32°C until colonies appeared. Plates were then replica plated on EMM-His and EMM-Ade to screen for colonies that would have undergone spontaneous recombination. Colonies that were His+Ade- were inoculated in liquid EMM-His and grown at 32°C for 24hs. After 24hs, cells were counted with a hemocytometer and released in 4mL YES at 100 cells/microliter. Released cells were placed at 32oC for 48hs. After 48hs, cells were counted again and plated onto EMM-Ade at 104-106 cells per plate. For each EMM-Ade plate we plated a YES control plate (1000cells/plate). Plates were incubated at 32°C for 4-5 days after which colonies on both YES and EMM-Ade plates were counted. The EMM-Ade plates were then replica plated onto EMM-Ade-His and re-incubated at 32°C for 3-4 more days. The colonies on EMM-Ade-His were then counted.

ura4-his3-ura4 recombination assay. Cells were struck onto YES from the -70°C freezer and grown at 32°C until colonies appeared. Cells were then inoculated into liquid EMM-His and grown at 32°C for 24h after which colonies were counted and released into 4mL YES at 100cells/microliter. Cultures were re- incubated at 32°C for 24hs after which they were once again counted and plated on EMM-Ura at 105-106 cells per plate. Because sometimes Ura- cells can cannibalize dead cells and may grow on media lacking uracil, Phloxin B was added to the media to eliminate false positives. Ura- colonies become dark red and are easily distinguished from the Ura+ ones. For each EMM-Ura plate a corresponding control YES plate was used at 1000cells/microliter. All plates were incubated at 32°C 4-5 days after which colonies were counted.

Data analysis. For all assays, the data were adjusted for viability and error in plating using the numbers on the YES plates. For example, if a YES plate produced 700 colonies it was assumed that either due to viability issues or miscalculation only 70% (or 0.7) viable cells produced colonies. If the corresponding EMM plate produced 100 colonies we adjusted this number by dividing 100/0.7 which equals 143. We assumed that had 100% of cells plated produced colonies we would have seen 143 colonies on EMM-his. This normalization was also important in order to control for systematic errors that might have been introduced as different people did the experiments.

For frequency graphs the data were analyzed and graphed using Excel. The graphs show the mean and standard error for each strain. For graphs showing rate the data was analyzed using the Lea-Coulson method (http://www.keshavsingh.org/protocols/FALCOR.html) and graphs were generated with the R statistical program. For the MHH recombination experiment data were computed as previously described (CULLEN *et al.* 2007).

SIMPLE Assay. *PCR amplification and cloning*. All PCR reactions were performed using the Phusion High-Fidelity DNA Polymerase (NEB), using the manufacturer's recommended specifications. After amplification, PCR products were first digested with DpnI (NEB) to eliminate the parental plasmid, cleaned up over columns (Qiagen), and then digested with either the SapI (NEB) or AarI (Thermo) restriction enzyme, following the manufacturer's specifications. After digestion, DNA products were cleaned up again as before, and 50 µg of the DNA was self-ligated using T4 DNA ligase (NEB). Ligations were transformed into chemo-competent Escherichia coli cells, and grown at 37°C overnight. Five colonies were screened from each cloning reaction by sequencing (GENEWIZ) or restriction digestion where applicable.

Plasmids. All site-specific mutations and deletions were performed on plasmid RP14 (*pUC19-ade6-M216*), which contains the *S. pombe ade6-M216* mutation (G46A) gene cloned in the Sal1 site.

Primers. The following primer pairs were used to generate the A46 → G mutation in ade6: (using the SapI enzyme): CGGGCTCTTCAGGCCGAATGATGGTA GAGGCAGCCCAT and GCGGCTCTTCTGCCCAA TTGACCACCTCCAAGGATCCCT; (using the AarI enzyme): CGGGCTCTTCATCGCGTGGAAATTA CGTTGTTCATCAACC and GCGGCTCTTCTCG AACCGTCGTAAGCCAATGTTTTACTTT. The fol- lowing primer pairs were used to generate the C1466 → T mutation in ADE6: (using the SapI enzyme): CGGGCTCTTCATTCGAGGTGTCCCTGT CGCCACTGTTGC and GCGGCTCTTCTGAAGCAT CTGAACAATAGAGTGAAGAG; (using the AarI enzyme): CGCCACCTGCGGATTTCGAGGTGT CCCTGTCGCCACTGTTGC and GCGCACCTG CCTGACGAAGCATCTGAACAATAGAGTGAAGAG. The following primer pairs were used to delete nucleotides 601–1101 in ADE6: (using the SapI enzyme): CGGGCTCTTCAGAATGTGAACGTAGG TATCAGATGCT and GCGGCTCTTCTTTCAA CTTTTCCGTCTAAACTGCGTACTA; (using the AarI enzyme): CGCCACCTGCGGATGAATGTGAAC GTAGGTATCAGATGCT and GCGCACCTGCCTG AATTCAACTTTTCCGTCTAAACTGCGTACT. The following primer pairs were used to insert the FLAG epitope tag after nucleotide 510 in ADE6: GCCGGC TCTTCTCGATGACGACAAGCTTGGTGATCG TCCGCTTTATGTTG and CGCCGCTCTTCATCGTC TTTGTAGTCTGCTTTGATGGCAGTAGGAATCTCA. The following primers were used for DNA sequencing: TAAAAACCTGTAAATGCTG, GATTCTGATTTAAG CAAG, and GGCCAAGAGAGTTTGGTTA.

CONCLUSION

Repair of DNA DSBs are very important in molecular biology. Methods for understanding the processes by which they are repaired can be powerful tools. There are many pathways that a DSB could be repaired by; and accurate determination of pathways is important in understanding the repair process. One way to study these aberrations is to introduce specific mutations at known loci. This can be accomplished via the SIMPLE method, allowing the researcher a great degree of control in their studies. Further, our labs Ura-His-Ura assay allows researchers to test specifically for single strand annealing. This assay also allows for homology on the same chromosome to be used in repair which would result in deletion of intervening sequences. We also showed that HIP1 biases recombination toward deletion via repair by SSA; and that cells lacking HIP1 have decreased recombination rates. These findings provide researchers with powerful tools for the further analysis of DNA damage repair, specifically that cells tend to bias repair towards deletion.

REFERENCES

Ahn, J. S., F. Osman and M. C. Whitby, 2005 Replication fork blockage by RTS1 at an ectopic site promotes recombination in fission yeast. EMBO J 24**:** 2011-2023.

Blackwell, C., K. A. Martin, A. Greenall, A. Pidoux, R. C. Allshire *et al.*, 2004 The Schizosaccharomyces pombe HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. Mol Cell Biol 24**:** 4309-4320.

Caldecott, K. W., 2008 Single-strand break repair and genetic disease. Nat Rev Genet 9**:** 619-631.

Cullen, J. K., S. P. Hussey, C. Walker, J. Prudden, B. Y. Wee *et al.*, 2007 Break-induced loss of heterozygosity in fission yeast: dual roles for homologous recombination in promoting translocations and preventing de novo telomere addition. Mol Cell Biol 27**:** 7745-7757.

DeGregori, J., 2013 Challenging the axiom: does the occurrence of oncogenic mutations truly limit cancer development with age? Oncogene 32**:** 1869-1875.

Gasparyan, H. J., J. Kroh, W. M. Michael and R. C. Petreaca, 2018 Development of the SapI/AarI Incision Mediated Plasmid Editing Method. J Mol Biol 430**:** 1426-1430.

Gomez, E. B., V. T. Angeles and S. L. Forsburg, 2005 A screen for Schizosaccharomyces pombe mutants defective in rereplication identifies new alleles of rad4+, cut9+ and psf2+. Genetics 169**:** 77-89.

Gomez, E. B., R. L. Nugent, S. Laria and S. L. Forsburg, 2008 Schizosaccharomyces pombe histone acetyltransferase Mst1 (KAT5) is an essential protein required for damage response and chromosome segregation. Genetics 179**:** 757-771.

Greenall, A., E. S. Williams, K. A. Martin, J. M. Palmer, J. Gray *et al.*, 2006 Hip3 interacts with the HIRA proteins Hip1 and Slm9 and is required for transcriptional silencing and accurate chromosome segregation. J Biol Chem 281**:** 8732-8739.

Grigaite, R., Z. Maneliene and A. Janulaitis, 2002 AarI, a restriction endonuclease from Arthrobacter aurescens SS2-322, which recognizes the novel non-palindromic sequence 5'-CACCTGC(N)4/8-3'. Nucleic Acids Res 30**:** e123.

Hastings, P. J., G. Ira and J. R. Lupski, 2009 A microhomology-mediated break-induced replication model for the origin of human copy number variation. PLoS Genet 5**:** e1000327.

Li, P. C., R. C. Petreaca, A. Jensen, J. P. Yuan, M. D. Green *et al.*, 2013 Replication fork stability is essential for the maintenance of centromere integrity in the absence of heterochromatin. Cell Rep 3**:** 638-645.

Ling, M. M., and B. H. Robinson, 1997 Approaches to DNA mutagenesis: an overview. Anal Biochem 254**:** 157-178.

Mehta, A., and J. E. Haber, 2014 Sources of DNA double-strand breaks and models of recombinational DNA repair. Cold Spring Harb Perspect Biol 6**:** a016428.

Meselson, M., and F. W. Stahl, 1958 The replication of DNA. Cold Spring Harb Symp Quant Biol 23**:** 9-12.

Osman, F., M. Adriance and S. McCready, 2000 The genetic control of spontaneous and UV-induced mitotic intrachromosomal recombination in the fission yeast Schizosaccharomyces pombe. Curr Genet 38**:** 113-125.

Osman, F., E. A. Fortunato and S. Subramani, 1996 Double-strand break-induced mitotic intrachromosomal recombination in the fission yeast Schizosaccharomyces pombe. Genetics 142**:** 341-357.

Osman, F., and M. C. Whitby, 2009 Monitoring homologous recombination following replication fork perturbation in the fission yeast Schizosaccharomyces pombe. Methods Mol Biol 521**:** 535-552.

Perry, G. H., N. J. Dominy, K. G. Claw, A. S. Lee, H. Fiegler *et al.*, 2007 Diet and the evolution of human amylase gene copy number variation. Nat Genet 39**:** 1256-1260.

Prudden, J., J. S. Evans, S. P. Hussey, B. Deans, P. O'Neill *et al.*, 2003 Pathway utilization in response to a site-specific DNA double-strand break in fission yeast. EMBO J 22**:** 1419-1430.

Ranjha, L., S. M. Howard and P. Cejka, 2018 Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. Chromosoma 127**:** 187-214.

Sallmyr, A., and A. E. Tomkinson, 2018 Repair of DNA double-strand breaks by mammalian alternative end-joining pathways. J Biol Chem 293**:** 10536-10546.

Tabancay, A. P., Jr., and S. L. Forsburg, 2006 Eukaryotic DNA replication in a chromatin context. Curr Top Dev Biol 76**:** 129-184.

Tessier, D. C., and D. Y. Thomas, 1993 PCR-assisted large insertion/deletion mutagenesis. Biotechniques 15**:** 498-501.

Vissers, L. E., S. S. Bhatt, I. M. Janssen, Z. Xia, S. R. Lalani *et al.*, 2009 Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. Hum Mol Genet 18**:** 3579-3593.

Watson, J. D., and F. H. Crick, 1974 Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. Nature 248**:** 765.

Yamaguchi-Iwai, Y., E. Sonoda, J. M. Buerstedde, O. Bezzubova, C. Morrison *et al.*, 1998 Homologous recombination, but not DNA repair, is reduced in vertebrate cells deficient in RAD52. Mol Cell Biol 18**:** 6430-6435.

Zheng, L., U. Baumann and J. L. Reymond, 2004 An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res 32**:** e115.